

A decomposition-based multiobjective evolutionary algorithm with angle-based adaptive penalty

Junfei Qiao ^{a,b,*}, Hongbiao Zhou ^{a,b}, Cuili Yang ^{a,b}, Shengxiang Yang ^{a,c}

^a College of Automation, Faculty of Information Technology, Beijing University of Technology, Beijing 100124, China

^b Beijing Key Laboratory of Computational Intelligence and Intelligent System, Beijing 100124, China

^c Centre for Computational Intelligence, School of Computer Science and Informatics, De Montfort University, Leicester LE1 9BH, UK

*Corresponding author at College of Automation, Faculty of Information Technology, Beijing University of Technology, Beijing 100124, China. Tel.: +86 13801263545

E-mail addresses: junfeiq@bjut.edu.cn (J.-F. Qiao), hyitzhb@163.com (H.-B. Zhou), 153003461@qq.com (C.-L. Yang), syang@dmu.ac.uk (S. Yang).

Abstract: A multiobjective evolutionary algorithm based on decomposition (MOEA/D) decomposes a multiobjective optimization problem (MOP) into a number of scalar optimization subproblems and optimizes them in a collaborative manner. In MOEA/D, decomposition mechanisms are used to push the population to approach the Pareto optimal front (POF), while a set of uniformly distributed weight vectors are applied to maintain the diversity of the population. Penalty-based boundary intersection (PBI) is one of the approaches used frequently in decomposition. In PBI, the penalty factor plays a crucial role in balancing convergence and diversity. However, the traditional PBI approach adopts a fixed penalty value, which will significantly degrade the performance of MOEA/D on some MOPs with complicated POFs. This paper proposes an angle-based adaptive penalty (AAP) scheme for MOEA/D, called MOEA/D-AAP, which can dynamically adjust the penalty value for each weight vector during the evolutionary process. Six newly designed benchmark MOPs and an MOP in the wastewater treatment process are used to test the effectiveness of the proposed MOEA/D-AAP. Comparison experiments demonstrate that the AAP scheme can significantly improve the performance of MOEA/D.

Keywords: multiobjective evolutionary algorithm; decomposition; penalty boundary intersection; angle-based adaptive penalty

1. Introduction

Multiobjective optimization problems (MOPs) which abide by the basic optimality principle of maximizing the benefit and minimizing the cost are dealt frequently in the industrial production and daily life [1]. Since these objectives usually conflict with each other, there does not exist a single solution that can optimize all the objectives simultaneously. For this sort of problems, Pareto optimal solutions may be a good choice for decision makers. Multiobjective evolutionary algorithms (MOEAs) can provide a set of representative Pareto solutions in a single run and are immune to the mathematical properties of objective functions. Hence, they have been attracted wide attention from researchers for solving MOPs [2].

There are three requirements in MOEAs: convergence, diversity, and coverage [3]. Convergence indicates that the distance between the solutions and the actual Pareto optimal front (POF) should be as small as possible. Diversity represents that the distribution of solutions along the actual POF should be as uniform as possible. Coverage denotes that the solutions should cover the entire POF. Selection, which is a core operation of MOEAs, is adopted to take care of above-mentioned three requirements [4]. According to the selection mechanism, current MOEAs can be roughly divided into three categories: dominance-based MOEAs, indicator-based MOEAs, and decomposition-

based MOEAs. In the first category, the Pareto dominance relationship is regarded as a main criterion to promote population convergence. At the same time, a density estimation strategy, such as crowding distance, is used to maintain population diversity. The typical representatives are MOPSO [5], NSGA-II [6], and SPEA2 [7]. In the second category, a performance indicator such as Hypervolume is adopted to guide the selection process. The typical representatives are SMS-EMOA [8] and HypE [9]. In the third category, a linear or nonlinear aggregation method, such as weighted sum, is applied to decompose a MOP into a number of single optimization subproblems. The typical representatives are MOGLS [10] and MOEA/D [11].

In the MOEA/D proposed by Zhang and Li [11], the selection of solutions depends on the aggregation function value, which greatly enhances the selection pressure toward the actual POF and gets more advantages in solving MOPs. At the same time, a set of uniformly distributed weight vectors are used to implicitly obtain good population diversity [12]. In addition, the concept of subproblem neighborhood, firstly presented in MOEA/D, is also able to improve the balance between exploration and exploitation of the algorithm in evolution. Since the original MOEA/D framework was proposed, many researchers have been trying to improve its performance from the following aspects.

- 1) Weight vector generation. Qi *et al.* [13] proposed a MOEA/D with adaptive weight adjustment (MOEA/D-AWA) scheme, which is more suitable for dealing with complex MOPs. Li *et al.* [14] designed a nonuniform mapping scheme by which the originally evenly distributed reference points on a canonical simplex can be mapped to new positions close to the decision maker-specified aspiration-level vector.
- 2) Decomposition method. Liu *et al.* [15] designed a new version of MOEA/D, termed MOEA/D-M2M, which decomposes a MOP into a set of simple multiobjective optimization subproblems. Experimental results show that the proposed MOEA/D-M2M is more competitive in solving MOPs. Wang *et al.* [16] proposed a MOEA/D with adaptive constrained decomposition (MOEA/D-ACD) approach, which can effectively enhance the diversity of nondominated solutions obtained. Cai *et al.* [17] integrated a constrained decomposition method with grid into MOEA/D, called CDG-MOEA, which can overcome the shortcomings of diversity loss in the existing decomposition approaches. Wu *et al.* [18] developed a MOEA/D with a learning-to-decompose paradigm (MOEA/D-LTD), which can adaptively set the decomposition method by periodically learning the characteristics of the estimated POF. Cai *et al.* [19] proposed a hybrid MOEA integrating domination-based sorting and decomposition (EAG-MOEA/D) for combinatorial optimization problems with two or three objectives. As an enhanced version of EAG-MOEA/D, Cai *et al.* [20] proposed a grid weighted sum Pareto local search (GWS-PLS) algorithm based on MOEA/D for combinatorial optimization problems.
- 3) Reproduction operation. Li *et al.* [21] designed a MOEA/D with adaptive operator selection (MOEA/D-FRRMAB) strategy to overcome the weaknesses caused by using simulated binary crossover (SBX) or differential evolution (DE) operator.
- 4) Mating selection and replacement mechanism. Li *et al.* [22] proposed a new version of MOEA/D with a simple and effective stable matching model, termed MOEA/D-STM, which can effectively balance the convergence and diversity during evolutionary search. Furthermore, an improved version of MOEA/D-STM, called MOEA/D-IR, is proposed in [23] for solving MOPs, which is designed based on the interrelationship selection scheme. Wang *et al.* [24] presented an adaptive replacement strategy for MOEA/D (MOEA/D-AGR), which can adjust the replacement neighborhood size dynamically.
- 5) Many-objective optimization. Li *et al.* [25] proposed a MOEA/D based on dominance and decomposition (MOEA/DD) to solve unconstrained and constrained many-objective optimization problems (MaOPs). Cai *et al.* [26] proposed a decomposition-based many-objective evolutionary algorithm with two types of adjustments for the direction vectors (MaOEA/D-2ADV) to solve MaOPs with irregular POFs. To address the problem of diversity estimation of POF approximations in the field of MaOPs, Cai *et al.* [27] proposed a unary diversity indicator based on reference vectors, which can provide either an offline or an online indicator for the selection of solutions.
- 6) Dynamic optimization. Chen *et al.* [28] implemented a dynamic two-archive evolutionary algorithm (DTAEA) on the original MOEA/D framework for solving dynamic multiobjective optimization problems (DMOPs) with a changing number of objectives.
- 7) Constrained optimization. Jan and Khanum [29] presented two penalty-parameterless constraint handling techniques and integrated them into MOEA/D to solve constrained MOPs. To balance convergence, diversity and feasibility simultaneously, Li *et al.* [30] proposed a parameter-free constrained handling technique, a two-archive evolutionary algorithm, for constrained multiobjective optimization. Since the decomposition approach has a significant influence on the performance of MOEA/D, we continue to conduct research along this direction in this paper.

The commonly used decomposition methods of MOEA/D are the weighted sum method, the Tchebycheff method,

and the penalty-based boundary intersection (PBI) method [31]. The weighted sum method has good performance in solving convex problems, but cannot effectively address concave problems. The Tchebycheff method can handle both convex and concave problems. If the number of weight vectors is not too large, the diversity of the approximated POF obtained by the PBI method is significantly better than that of the Tchebycheff method [32]. In the PBI method, a scalar optimization subproblem associated with a weight vector is constructed by a perpendicular distance and a utopia distance through a penalty factor θ . And then all the subproblems are optimized in a collaborative manner for approaching the actual POF. Therefore, the penalty factor in the PBI method has a significant impact on the performance of MOEA/D [33].

In the original PBI method, a fixed θ value which remains constant in the search process is adopted. However, it is difficult to obtain better distributed solutions for MOEA/D with fixed θ value on MOPs with complex POFs (e.g., discontinuous POFs or POFs with a sharp peak and long tail) and practical engineering optimization problems. But, due to the lack of testing instances with complicated characteristics (e.g., the ZDT, DTLZ, and WFG benchmark test suits, often used in MOEAs, do not have complex POFs), there is relatively little research on the PBI penalty factor. In [34], Sato proposed an inversed PBI method, which improves the search ability of the algorithm. However, this method still does not consider the influence of the penalty factor, and it is difficult to get the balance between convergence and diversity when solving MOPs with irregular POFs. In [35], Yang et al. presented an adaptive penalty scheme (APS) where the penalty factor θ linearly increases with the iteration number. In the APS strategy, a small penalty factor is used at the early stage to accelerate the convergence speed. On the contrary, a large penalty factor is adopted at the late stage to obtain more uniformly distributed solutions. The experimental results show that the APS scheme can obtain better distributed approximated POFs in solving MOPs with complex characteristics. However, this method does not consider the evolution state of each subproblem and the unified θ value is not suitable for maintaining the diversity of the boundary solutions.

To solve the above-mentioned problem, an angle-based adaptive penalty (AAP) scheme, based on the angle information between solutions and weight vectors, is designed in this paper and integrated into MOEA/D, called MOEA/D-AAP. The effectiveness of the proposed MOEA/D-AAP algorithm is verified based on six benchmark MOPs with complex POFs and an MOP in the wastewater treatment process (WWTP).

The rest of this paper is organized as follows. Section 2 introduces some preliminaries on multiobjective optimization and the PBI decomposition method. In Section 3, the influence of the penalty factor is investigated, and then the AAP scheme is described. In Section 4, the AAP scheme is integrated into MOEA/D. Section 5 introduces the benchmark test functions with complex POFs and the performance metrics used in this paper. The experimental studies and discussion are presented in Section 6. Finally, Section 7 concludes this paper.

2. Preliminaries

This section first gives some basic definitions of multiobjective optimization. Then, the PBI decomposition method is introduced briefly.

2.1. Basic definitions

Without loss of generality, we consider minimization problems as objective functions in this paper. Then, an MOP can be described as follows:

$$\begin{aligned} & \text{minimize} \quad \mathbf{F}(\mathbf{x}) = (f_1(\mathbf{x}), f_2(\mathbf{x}), \dots, f_m(\mathbf{x}))^T \\ & \text{subject to} \quad \mathbf{x} \in \Omega \end{aligned} \quad (1)$$

where $\mathbf{x}=(x_1, \dots, x_n)^T \in \Omega$ is a decision variable vector, Ω is the feasible search region, $\mathbf{F}: \Omega \rightarrow R^m$ consists of m real-valued objective functions, and R^m is the objective space [11].

Let $\mathbf{x}, \mathbf{y} \in \Omega$ are two feasible solutions of an MOP as defined in Eq. (1), \mathbf{x} is said to dominate \mathbf{y} , denoted by $\mathbf{x} \prec \mathbf{y}$, if and only if $f_i(\mathbf{x}) \leq f_i(\mathbf{y})$ for every $i \in \{1, \dots, m\}$ and $f_j(\mathbf{x}) < f_j(\mathbf{y})$ for at least one index $j \in \{1, \dots, m\}$. A point $\mathbf{x}^* \in \Omega$ is Pareto-optimal to the MOP defined in Eq. (1) if there exists no point $\mathbf{x} \in \Omega$ such that $\mathbf{F}(\mathbf{x})$ dominates $\mathbf{F}(\mathbf{x}^*)$. $\mathbf{F}(\mathbf{x}^*)$ is then called a Pareto-optimal (objective) vector. In other words, any improvement in one objective of a Pareto-optimal solution must lead to deterioration in at least one other objective. The set of all the Pareto-optimal points is

called the Pareto-optimal set (POS) and the set of all the Pareto-optimal objective vectors is called the POF [22].

2.2. The PBI decomposition method

In this study, the decomposition method which be used in MOEA/D is the PBI method. In the PBI method, the scalar objective optimization problem is in the form

$$\begin{aligned} & \text{minimize } g^{pbi}(\mathbf{x}|\mathbf{w}, \mathbf{z}^*) = d_1 + \theta d_2 \\ & \text{subject to } \mathbf{x} \in \Omega \end{aligned} \quad (2)$$

$$\text{where } d_1 = \frac{\|(\mathbf{F}(\mathbf{x}) - \mathbf{z}^*)^T \mathbf{w}\|}{\|\mathbf{w}\|} \text{ and } d_2 = \left\| \mathbf{F}(\mathbf{x}) - \left(\mathbf{z}^* + d_1 \frac{\mathbf{w}}{\|\mathbf{w}\|} \right) \right\|.$$

$\theta > 0$ is a predefined penalty parameter, $\mathbf{w} = (w_1, \dots, w_m)^T$ is the search direction vector of the subproblem, $\mathbf{z}^* = (z_1^*, \dots, z_m^*)$ is the ideal point in the objective space, i.e., $z_i^* = \min \{f_i(\mathbf{x}) | \mathbf{x} \in \Omega\}$ for each $i = 1, \dots, m$. Since \mathbf{z}^* is generally unknown before searching, z_i^* can be replaced by $f_i(\mathbf{x})$ with the smallest value during the searching process [22], [23].

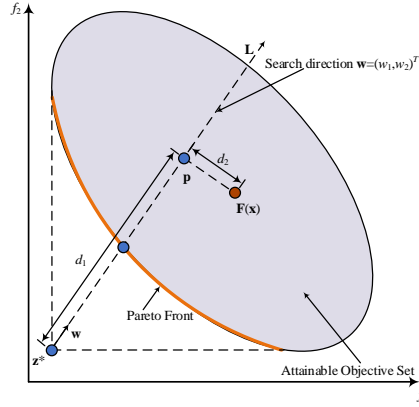


Fig. 1 Illustration of PBI

Fig. 1 gives a brief illustration of the PBI decomposition method, where \mathbf{L} is a straight line passing through \mathbf{z}^* with direction \mathbf{w} , \mathbf{p} is the projection of $\mathbf{F}(\mathbf{x})$ on \mathbf{L} , d_1 is the distance between \mathbf{z}^* and \mathbf{p} , and d_2 is the vertical distance between $\mathbf{F}(\mathbf{x})$ and \mathbf{L} . The purpose of this method is to push $\mathbf{F}(\mathbf{x})$ as low as possible by minimizing the function g^{pbi} , so that it can approximate the edge of the attainable objective set. It is not difficult to find that θ plays a key role in balancing convergence (represented by d_1) and diversity (represented by d_2). A small value of θ can accelerate the convergence speed of the algorithm, but it may induce the population to be trapped into a local POF. On the other hand, a large value of θ is helpful for maintaining the diversity of population, but it will slow down the convergence process [25], [35].

3. Angle-based adaptive penalty scheme

3.1. Influence of the penalty factor

This section further elaborates the influence of the value of θ on the search performance of the MOEA/D with PBI, so as to summarize the law and introduce our adaptive penalty scheme. As discussed in [35], Fig. 2 shows the impact of the penalty factor on the search performance of the algorithm, where \mathbf{w}^1 and \mathbf{w}^2 are two weight vectors, points A to F are distributed solutions along the POF, and G and H are two solutions in the objective space. θ is equal to 1.0 and 5.0 in Fig. 2 (a) and Fig. 2 (b), respectively.

In Fig. 2 (a), point B of the subproblem associated with \mathbf{w}^1 will be replaced by point C when the algorithm performs the replacement operation, since $g^{pbi}(\mathbf{C}|\mathbf{w}^1, \mathbf{z}^*)$ is less than $g^{pbi}(\mathbf{B}|\mathbf{w}^1, \mathbf{z}^*)$. As a result, solutions tend to be concentrated in the middle region of the POF, reducing the diversity of the population. In other words, to enhance

the diversity, point C should not be replaced by point B during the evolutionary process. However, if the penalty factor associated with \mathbf{w}^1 is too small, it is obvious that this is difficult to achieve. Therefore, in this case, the penalty factor value should be increased appropriately so that point C cannot be replaced by point B, thus maintaining the population diversity.

In Fig. 2 (b), because $g^{pbi}(H|\mathbf{w}^1, \mathbf{z}^*)$ is larger than $g^{pbi}(G|\mathbf{w}^1, \mathbf{z}^*)$, point G cannot be replaced by point H, thus resulting in the reduction of the convergence speed. In other words, to speed up the convergence speed, point G should be replaced by point H during the evolutionary process. However, if the penalty factor associated with \mathbf{w}^1 is too big, it is obvious that this is difficult to achieve. Therefore, in this case, the penalty factor value should be decreased appropriately so that point G can be replaced by point H, thus promoting the population convergence.

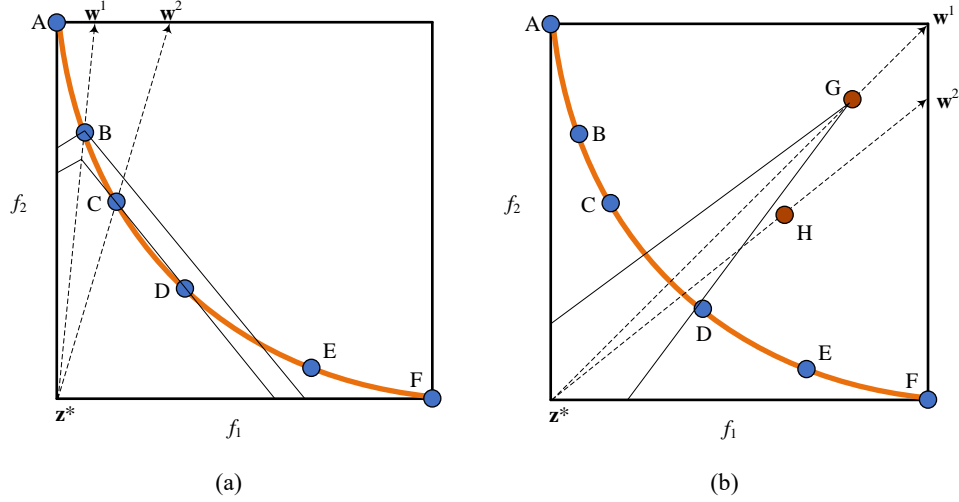


Fig. 2 Influence of the penalty factor on search performance
(a) $\theta = 1.0$; (b) $\theta = 5.0$

3.2. Advantages of the adaptive PBI

Fig. 3 shows the advantages of the adaptive PBI strategy. From Fig. 3(a), for the plain PBI method, the population tends to converge to the middle region of the true POF, as the value of the penalty factor θ is fixed (e. g., θ is set to 5 in the original MOEA/D). If the value of θ is increased, the evolutionary speed will slow down, so that there may be no guarantee for the convergence of the approximated POF within a limited number of iterations. From Fig. 3(b), for the adaptive PBI method, since different subproblems on different evolutionary stages may have different θ values, both the uniformity of the approximated POF and the quickness of the evolutionary process can be strengthened.

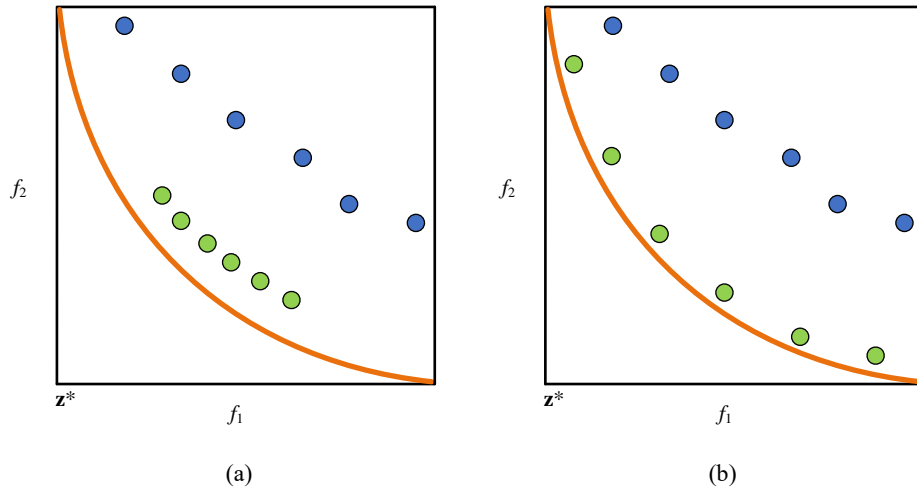


Fig. 3 The effectiveness of the adaptive PBI with respect to the plain PBI
(a) plain PBI; (b) adaptive PBI

3.3. Angle-based adaptive penalty scheme

In order to adaptively adjust the corresponding θ value of each subproblem according to the evolutionary state, we first define the solution concentration of a weight vector, and the specific formula is as follows:

$$con^i = \frac{nc^i - nc_{\min} + \varepsilon}{nc_{\max} - nc_{\min} + \varepsilon}, i = 1, \dots, N \quad (3)$$

where $con^i \in [0,1]$ is the solution concentration of the weight vector \mathbf{w}^i , nc^i is the number of solutions gathered around the weight vector \mathbf{w}^i . nc_{\min} and nc_{\max} are the minimum and maximum value of $nc(W)$, respectively. If the acute angle between solution \mathbf{x}^j and weight vector \mathbf{w}^i is less than that of the maximum acute angle between weight vectors in neighborhood, it can be considered that \mathbf{x}^j is gathered around \mathbf{w}^i . ε is set to be 10^{-6} to prevent a zero denominator. nc^i is calculated by

$$nc^i = \sum_{j=1}^N \gamma^{ij}, i = 1, \dots, N \quad (4)$$

where γ^{ij} is used to indicate whether solution \mathbf{x}^j is gathered around weight vector \mathbf{w}^i or not. If \mathbf{x}^j is gathered around \mathbf{w}^i , $\gamma^{ij}=1$; otherwise, $\gamma^{ij}=0$. γ^{ij} is defined as follows:

$$\gamma^{ij} = \begin{cases} 1, & \text{if } \beta^{ij} < \beta_{\text{th}} \\ 0, & \text{otherwise} \end{cases} \quad (5)$$

However, it is worth noting that, if the POF of the MOP in question has a sharp peak and long tail, the acute angle between a solution and the weight vector passed through the ideal point \mathbf{z}^* may have a very small value. Thus, the solution which is far from the weight vector may be improperly regarded as one gathered around the weight vector. Therefore, in this study, the acute angle is calculated by the solution and the weight vector passed through the reference point \mathbf{z}^r . β^{ij} is defined as

$$\beta^{ij} = \arccos \left(\frac{(\mathbf{z}^r - \mathbf{F}(\mathbf{x}^j))^T (\mathbf{z}^r - \mathbf{w}^i)}{\|\mathbf{z}^r - \mathbf{F}(\mathbf{x}^j)\| \|\mathbf{z}^r - \mathbf{w}^i\|} \right) \quad (6)$$

$i = 1, \dots, N, j = 1, \dots, N$

For the sake of simplicity, the value of β_{th} is set as the average of all angles between any two adjacent weight vectors multiplied by T , as follows:

$$\beta_{\text{th}} = \frac{T}{N-1} \sum_{i=1}^{N-1} \beta^i \quad (7)$$

where T is the neighborhood size, and β^i is the acute angle between $\mathbf{z}^r - \mathbf{w}^i$ and $\mathbf{z}^r - \mathbf{w}^{i+1}$, as follows:

$$\beta^i = \arccos \left(\frac{(\mathbf{z}^r - \mathbf{w}^i)^T (\mathbf{z}^r - \mathbf{w}^{i+1})}{\|\mathbf{z}^r - \mathbf{w}^i\| \|\mathbf{z}^r - \mathbf{w}^{i+1}\|} \right) \quad (8)$$

$i = 1, \dots, N-1$

In order to illustrate the advantages of the above method, Fig. 4 gives a schematic diagram for calculating the acute angle β , in which \mathbf{w}^1 and \mathbf{w}^2 are two weight vectors. From the view of ideal point \mathbf{z}^* , the angle α between point A and \mathbf{w}^2 is less than the angle α_{th} between \mathbf{w}^1 and \mathbf{w}^2 , so point A will be judged to be clustered around the weight vector \mathbf{w}^2 . From Figure 4, point B is more likely to be awarded to \mathbf{w}^2 . In fact, there is no solution around \mathbf{w}^2 . In the evolution process, we need to increase the penalty factor of the subproblem associated with the weight vector \mathbf{w}^2 to enhance the diversity of boundary solutions. From the perspective of reference point \mathbf{z}^r , the angle β between point A and \mathbf{w}^2 is larger than the angle β_{th} between \mathbf{w}^1 and \mathbf{w}^2 , so point A will not be awarded to \mathbf{w}^2 .

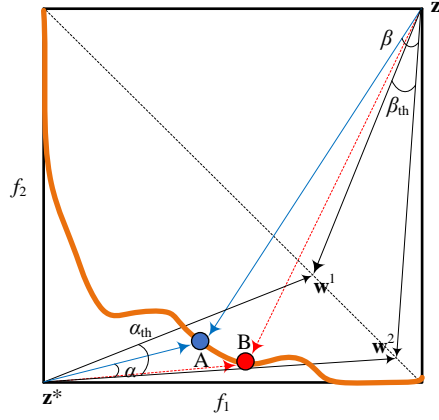


Fig. 4 Schematic diagram for calculating acute angle β

Since the optimal solution to a subproblem should be located around the search direction vector, if the number of solutions surrounding some direction vectors is larger than that of other direction vectors, then it means that the solutions are not evenly distributed. In the evolutionary process, if the penalty parameter is too small, the solution to the subproblem associated with a boundary weight vector will be easily replaced by the solution with better convergence in the neighborhood, which leads to the loss of diversity; on the other hand, when the penalty parameter is too large, the convergence speed will slow down, leading to the failure of approaching to the true POF. Therefore, the penalty factor of each subproblem should be dynamically adjusted according to the solution concentration of each weight vector.

Generally, linear method, exponential method and Sigmoid function method are the three most commonly used representation methods [24]. For the linear method, the change of θ values in the boundary region is too drastic, which may lead to a loss diversity of the boundary solutions. For the exponential method, the change of θ values in the middle region is too slow, which is likely to affect the replacement operation among elite solutions. The Sigmoid function method can overcome these shortcomings. Therefore, in this study, the Sigmoid function method is adopted to characterize the relationship between the penalty factor θ of a subproblem and the concentration con , as follows:

$$\theta(con) = \frac{1}{1 + ae^{-b \times con}} \quad (9)$$

when con changes from 0 to 1, θ changes from $\zeta \times N$ to 2. ζ is a positive value controlling the magnitude of θ and $\zeta=0.5$ is recommended in this study. That is, when $con=0$, $\theta=50$ and when $con=1$, $\theta=2$. The analysis on how to set the range of θ is presented in Section 6.2. Based on Eq. (9), we can get $a=-0.98$ and $b=0.6729$, so the relationship between θ^i of the i th subproblem and its solution concentration con^i can be obtained as

$$\theta^i = \frac{1}{1 - 0.98e^{-0.6729con^i}} \quad (10)$$

According to Eq. (10), the balance between convergence and diversity can be achieved by dynamically adjusting the penalty factor of each subproblem in the evolution. In the plain PBI-based MOEA/D, the boundary solutions can easily gather around the middle subproblems for the test instances with complicated POFs, which leads to the diversity loss of boundary solutions. However, in the adaptive PBI-based MOEA/D, for a boundary subproblem, a larger penalty factor will help to enhance the population diversity when it has a low solution concentration; for a middle subproblem, a smaller penalty factor will help to maintain the population convergence when it has a high solution concentration.

4. The MOEA/D-AAP Algorithm

The proposed MOEA/D-AAP is developed based on MOEA/D-DE [12], which is a well-known improved version of MOEA/D. Here, we first briefly introduce the initialization and reproduction operations, and then give the procedure of MOEA/D-AAP algorithm.

4.1. Initialization

In MOEA/D, the initialization operation mainly includes the initialization of population and weight vectors. Since we have no prior knowledge about the position of the POS, the initial population is randomly generated from the decision space. During evolution, the optimization of each subproblem is completed by evolutionary operation between the subproblem and subproblems in its neighborhood. The relationship of adjacent subproblems is determined by the distance between weight vectors associated with subproblems. To a certain extent, the uniform distribution of weight vectors can improve the uniformity of the approximated POF obtained by the algorithm. MOEA/D adopts the simplex lattice design method, proposed by Scheff in 1958, to set weight vectors, as follows [22]:

$$w_j^i \in \left\{ \frac{0}{H}, \frac{1}{H}, \frac{2}{H}, \dots, \frac{H}{H} \right\} \quad (11)$$

$$\sum_{j=1}^m w_j^i = 1 \quad (12)$$

where each subproblem $i(i=1,2,\dots,N)$ corresponds to a weight vector $\mathbf{w}^i=(w_1^i, w_2^i, \dots, w_m^i)$, $\mathbf{w}_j^i \geq 0$, $j=1,2,\dots,m$, the weight vector set is $W=(\mathbf{w}^1, \mathbf{w}^2, \dots, \mathbf{w}^N)$, where $N = C_{H+m-1}^{m-1}$ is the total number of weight vectors.

4.2. Reproduction Operation

Similar to MOEA/D-DE [12], MOEA/D-AAP also uses differential evolution operation and polynomial mutation to produce an offspring population $Q_i=\{\mathbf{x}'^1, \mathbf{x}'^2, \dots, \mathbf{x}'^N\}$, where the offspring solution $\mathbf{x}'^i=\{x_1'^i, x_2'^i, \dots, x_n'^i\}$ is produced as follows:

$$u_j^i = \begin{cases} x_j^{r_1} + F \times (x_j^{r_2} - x_j^{r_3}) & \text{if rand} < CR \\ x_j^i & \text{otherwise} \end{cases} \quad (13)$$

where $j \in \{1,2,\dots,n\}$, $\text{rand} \in [0,1]$, F and CR are two control parameters. Then, the polynomial mutation operation is performed as

$$x_j'^i = \begin{cases} u_j^i + \sigma_j \times (b_j - a_j) & \text{if rand} < p_m \\ u_j^i & \text{otherwise} \end{cases} \quad (14)$$

$$\sigma_j = \begin{cases} (2 \times \text{rand})^{\frac{1}{\eta+1}} - 1 & \text{if rand} < 0.5 \\ 1 - (2 - 2 \times \text{rand})^{\frac{1}{\eta+1}} & \text{otherwise} \end{cases} \quad (15)$$

where p_m is the mutation probability, η is the distribution index, a_j and b_j are the lower and upper bounds of the j th decision variable, respectively.

4.3. MOEA/D-AAP algorithm procedure

Table 1 shows the procedure of the MOEA/D-AAP algorithm.

Table 1 The procedure of MOEA/D-AAP

| Algorithm: MOEA/D-AAP | |
|-----------------------|---|
| 1 | Initialize the population $P \leftarrow \{\mathbf{x}^1, \dots, \mathbf{x}^N\}$, a set of weight vector $W \leftarrow \{\mathbf{w}^1, \dots, \mathbf{w}^N\}$, the ideal and reference objective vectors $\mathbf{z}^*, \mathbf{z}^r$; |
| 2 | Set $iteration \leftarrow 0$; |
| 3 | for $i \leftarrow 1$ to N do |
| 4 | $B(i) \leftarrow \{i_1, \dots, i_T\}$ where $\mathbf{w}^{i_1}, \dots, \mathbf{w}^{i_T}$ are the T closest weight vectors to \mathbf{w}^i ; |
| 5 | end |
| 6 | Calculate the threshold of acute angle β_{th} by Eqs. (7) and (8); |
| 7 | while <i>Stopping criterion is not satisfied</i> do |
| 8 | for $i \leftarrow 1$ to N do |
| 9 | for $j \leftarrow 1$ to N do |

```

10     Calculate the acute angle  $\beta^j$  between weight vector  $\mathbf{w}^i$  and solution  $\mathbf{x}^j$  by Eq. (6);
11     end
12 end
13     Calculate the number of solutions  $nc$  around each weight vector by Eqs. (4) and (5);
14     Calculate the concentration of solutions  $con$  for each weight vector by Eq. (3);
15     Adjust the value of  $\theta$  for each weight vector by Eq. (10);
16     for  $i \leftarrow 1$  to  $N$  do
17         if  $uniform(0,1) < \delta$  then
18              $E \leftarrow B(i)$ ;
19         else
20              $E \leftarrow P$ ;
21         end
22         Randomly select three solutions  $\mathbf{x}^{*1}$ ,  $\mathbf{x}^{*2}$ , and  $\mathbf{x}^{*3}$  from  $E$ ;
23         Generate a candidate  $\mathbf{x}'$  by using the method described in section 4.2;
24         Evaluate the F-function value  $FV$  of  $\mathbf{x}'$ ;
25         Update the current ideal objective vector  $\mathbf{z}^*$ ;
26         Set  $c = 0$ . while  $c < nr$  and  $E \neq \emptyset$  do
27             Randomly pick an index  $j$  from  $E$ ;
28             Remove  $j$  from  $E$ ;
29             Calculate the  $g$  value using PBI method with  $\theta^j$  for  $\mathbf{w}^j$ ;
30             if  $g(\mathbf{x}'|\mathbf{w}^j, \mathbf{z}^*) \leq g(\mathbf{x}^j|\mathbf{w}^j, \mathbf{z}^*)$  then
31                 Set  $\mathbf{x}^j = \mathbf{x}'$ ,  $FV^j = FV(\mathbf{x}')$ , and  $c = c + 1$ ;
32             end
33         end
34     end
35      $iteration++$ ;
36 end
37 return  $P$ ;

```

4.4. Discussion

In the following, the main idea of the proposed MOEA/D-AAP is compared with that of the MOEA/D-ACD [16], CDG-MOEA [17], and MOEA/D-APS [35].

In MOEA/D-ACD, the constrained subproblem is in the form:

$$\begin{aligned}
 & \text{minimize } g^{\text{pbi}}(\mathbf{x}|\mathbf{w}, \mathbf{z}^*) \\
 & \text{subject to } \langle \mathbf{w}^i, \mathbf{F}(\mathbf{x}) - \mathbf{z}^* \rangle \leq 0.5\alpha^i \\
 & \mathbf{x} \in \Omega
 \end{aligned} \tag{16}$$

where α^i is a control parameter for defining the improvement region. $\langle \mathbf{w}^i, \mathbf{F}(\mathbf{x}) - \mathbf{z}^* \rangle$ is the acute angle between \mathbf{w}^i and $\mathbf{F}(\mathbf{x}) - \mathbf{z}^*$, as shown in Fig. 5 (a). When performing replacement operation, MOEA/D-ACD should consider not only the aggregation function value of subproblem, but also the inequality constraint defined by the Eq. (16). In MOEA/D-ACD, a strategy for adaptively adjusting the value of α^i is designed for improving the population diversity. The performance of MOEA/D-ACD is validated on ZDT and MOP test suits. In MOEA/D-AAP, the distribution of the current solutions is assessed by the solution concentration of the weight vector, and the acute angle between solution and weight vector is utilized to design an adaptive adjustment strategy for penalty factor. At first glance, it is a little similar to MOEA/D-ACD. However, it can be clearly seen from Fig. 4 that the angle calculation method designed in MOEA/D-AAP adopts the nadir point as reference point to solve complicated MOPs with a sharp peak and long tail. In such a way, more boundary solutions can be found at different search stage, which can improve the performance of MOEA/D significantly. Probably due to the lack of targeted design, the performance of MOEA/D-ACD is a little less than that of the proposed MOEA/D-AAP on F1-F6.

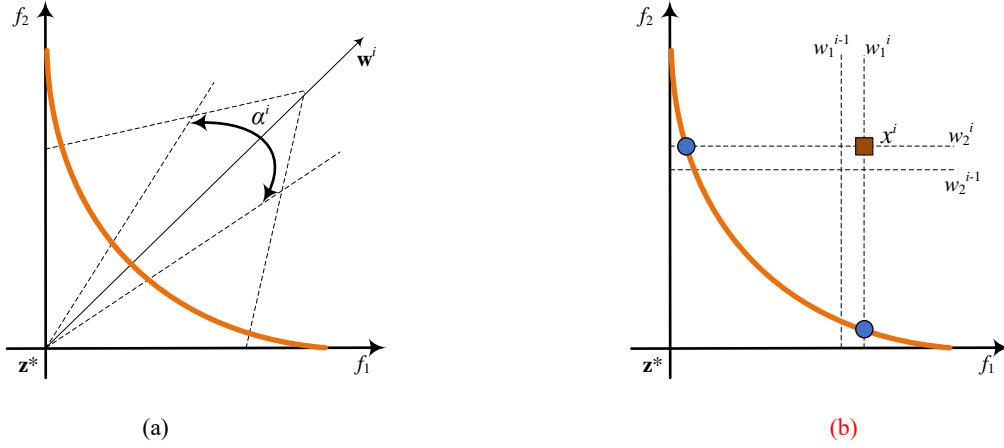


Fig. 5 Illustration of the improvement regions

(a) constrained PBI method; (b) constrained decomposition with grids

As an important milestone study, CDG-MOEA redefines the improvement region of classical decomposition method by using constrained decomposition with grids (CDG) strategy to prevent the same solution from being assigned to many different subproblems, thus maintaining the diversity of the population. Fig. 5 (b) shows the improvement regions of CDG. It can be observed that the volumes of the improvement regions are significantly reduced to the narrowed regions, where the same solution can be assigned to at most m subproblems. By using CDG strategy, CDG-MOEA can effectively handle MOPs with complicated POFs. However, the core idea of the proposed MOEA/D-AAP in this paper is quite different from that of the CDG-MOEA. MOEA/D-AAP can obtain satisfactory approximation results without destroying the basic framework of the original MOEA/D algorithm. It is easy for domain engineers to solve practical engineering problems by using MOEA/D-AAP.

In MOEA/D-APS, the adjustment approach of θ is defined as follows:

$$\theta = \theta_{\min} + (\theta_{\max} - \theta_{\min}) \frac{t}{\max\text{Iteration}} \quad (17)$$

where $\theta_{\min}=1.0$, $\theta_{\max}=10.0$, and $\max\text{Iteration}$ is the maximum number of iterations. From Eq. (17), with the increase of the iteration number, the value of θ will linearly increase. At the early stage, a small value of θ can promote the convergence speed; at the late stage, a large value of θ can improve the solution diversity. However, in the APS method, each subproblem has the same penalty value in the evolutionary process. In practical, the boundary subproblems and the intermediate subproblems should be treated differently for obtaining a well distributed approximated POF.

4.5. Computational cost of MOEA/D-AAP

In this section, let us consider the computational cost of MOEA/D-AAP in one generation. MOEA/D-AAP has the same framework as MOEA/D-DE [12], thus the increased computation cost is attributed to its detection step for the AAP scheme. The calculation of the threshold of acute angle β_{th} (line 6 in Table 1) requires $O(mN)$ computations. During the main while loop, the calculation of the acute angle β^{ij} (line 8 to line 12 in Table 1) costs $O(mN^2)$ computations. Next, the calculation of the solution concentration con requires $O(N^2 \log N)$ comparisons. Therefore, the overall complexity of AAP strategy in one generation is $O(N^2 \log N)$.

5. Experimental setting

5.1. Test problems

In order to evaluate the effectiveness of the proposed MOEA/D-AAP, six benchmark test instances with complex POFs (F1-F6) and the MOP in the WWTP are considered in our experimental studies. The detailed description of

these benchmark MOPs (F1-F6) is presented in Table 2 [35], and the MOP in the WWTP is described in the Section 6.2.

Table 2 Test instances

| Instance | Description | Domain | Number of variables | Notes |
|----------|--|------------|---------------------|--------------------------|
| F1 | $f_1(x) = (1 + g(x))x_1$ $f_2(x) = (1 + g(x))(1 - \sqrt{x_1})^3$ $g(x) = 2\sin(0.5\pi x_1)(n - 1 + \sum_{i=2}^n (y_i^2 - \cos(2\pi y_i)))$ where $y_{i=2n} = x_i - \sin(0.5\pi x_i)$ POF: $f_2 = (1 - \sqrt{f_1})^3$ POS: $x_i = \sin(0.5\pi x_i), i = 2, \dots, n$ | $[0, 1]^n$ | 20 | Uni-modal convex |
| F2 | $f_1(x) = (1 + g(x))x_1$ $f_2(x) = (1 + g(x))\sqrt{1 - x_1^5}$ $g(x) = 2\sin(0.5\pi x_1)(n - 1 + \sum_{i=2}^n (y_i^2 - \cos(2\pi y_i)))$ where $y_{i=2n} = x_i - \sin(0.5\pi x_i)$ POF: $f_2 = \sqrt{1 - f_1^5}$ POS: $x_i = \sin(0.5\pi x_i), i = 2, \dots, n$ | $[0, 1]^n$ | 20 | Uni-modal concave |
| F3 | $f_1(x) = (1 + g(x))x_1$ $f_2(x) = \frac{1}{2}(1 + g(x))(1 - x_1^{0.1} + (1 - \sqrt{x_1})^2 \cos^2(3\pi x_1))$ $g(x) = 2\sin(0.5\pi x_1)(n - 1 + \sum_{i=2}^n (y_i^2 - \cos(2\pi y_i)))$ where $y_{i=2n} = x_i - \sin(0.5\pi x_i)$ POF: $f_2 = \frac{1}{2}(1 - f_1^{0.1} + (1 - \sqrt{f_1})^2 \cos^2(3\pi f_1))$ POS: $x_i = \sin(0.5\pi x_i), i = 2, \dots, n$ | $[0, 1]^n$ | 20 | Multi-modal disconnected |
| F4 | $f_1(x) = (1 + g(x))(x_1 + 0.05\sin(6\pi x_1))^2$ $f_2(x) = (1 + g(x))(1 - x_1 + 0.05\sin(6\pi x_1))^2$ $g(x) = 2\sin(0.5\pi x_1)(n - 1 + \sum_{i=2}^n (y_i^2 - \cos(2\pi y_i)))$ where $y_{i=2n} = x_i - \sin(0.5\pi x_i)$ POF: $f_1^{0.5} + f_2^{0.5} = 1 + 0.1\sin(3\pi(f_1^{0.5} - f_2^{0.5} + 1))$ POS: $x_i = \sin(0.5\pi x_i), i = 2, \dots, n$ | $[0, 1]^n$ | 20 | Multi-modal mixed |
| F5 | $f_1(x) = (1 + g(x))(x_1 + 0.05\sin(6\pi x_1))^{0.2}$ $f_2(x) = (1 + g(x))(1 - x_1 + 0.05\sin(6\pi x_1))^{10}$ $g(x) = 2\sin(0.5\pi x_1)(n - 1 + \sum_{i=2}^n (y_i^2 - \cos(2\pi y_i)))$ where $y_{i=2n} = x_i - \sin(0.5\pi x_i)$ POF: $f_1^5 + f_2^{0.1} = 1 + 0.1\sin(3\pi(f_1^5 - f_2^{0.1} + 1))$ POS: $x_i = \sin(0.5\pi x_i), i = 2, \dots, n$ | $[0, 1]^n$ | 20 | Multi-modal mixed |
| F6 | $f_1(x) = ((1 + g(x))\cos(0.5\pi x_1)\cos(0.5\pi x_2))^4$ $f_2(x) = ((1 + g(x))\cos(0.5\pi x_1)\sin(0.5\pi x_2))^4$ $f_3(x) = ((1 + g(x))\sin(0.5\pi x_1))^2$ $g(x) = \sum_{i=3}^n (x_i - 0.5)^2$ POF: $\sqrt{f_1} + \sqrt{f_2} + f_3 = 1$ POS: $x_i = 0.5, i = 3, \dots, n$ | $[0, 1]^n$ | 20 | Uni-modal convex |

5.2. Performance metrics

In our empirical studies, we consider the following **three** widely used performance metrics [2].

1) Inverted Generational Distance (IGD)

Let S^* be a set of points uniformly sampled from the true POF, and S be the set of approximated solutions obtained

by an MOEA, the IGD indicator measures the gap between S^* and S , which is calculated as follows:

$$IGD(S^*, S) = \frac{\sum_{\mathbf{x} \in S^*} d(\mathbf{x}, S)}{|S^*|} \quad (18)$$

where $d(\mathbf{x}, S)$ is the Euclidean distance between the solution \mathbf{x} and its nearest neighbor in S , and $|S^*|$ is the cardinality of S^* . If the number of points in S^* is big enough, the IGD indicator can measure the convergence and diversity of the approximated POF obtained by an MOEA at the same time. The smaller the IGD value, the better the quality of the approximated POF. Since the true POF of the test instance needs to be known in advance when calculating the IGD, the IGD indicator is only used for F1-F6 test instances. In the experimental studies, 500 uniformly distributed points are sampled from the true POF for bi-objective test instances, and 1000 for three-objective ones, respectively.

2) Hypervolume (HV)

Let $\mathbf{z}^r = (z_1^r, z_2^r, \dots, z_m^r)^T$ be a reference point in the objective space that is dominated by all points on the true POF, and S be the set of approximated solutions obtained by an MOEA, the HV indicator measures the size of objective space dominated by the solution in S and bounded by \mathbf{z}^r

$$HV(S) = \text{VOL}(\bigcup_{\mathbf{x} \in S} [f_1(\mathbf{x}), z_1^r] \times \dots \times [f_m(\mathbf{x}), z_m^r]) \quad (19)$$

where VOL represents the Lebesgue measure. The bigger the HV value, the better the quality of the approximated POF. In the experimental studies, we set $\mathbf{z}^r = (1.1, 1.1)^T$ for bi-objective test instances, $\mathbf{z}^r = (1.2, 1.2)^T$ for the MOP in the WWTP, and $\mathbf{z}^r = (1.1, 1.1, 1.1)^T$ for three-objective test instances, respectively.

3) Δ Metric

Δ Metric [6, 27] measures the diversity of the approximation S as follows:

$$\Delta = \frac{d_f + d_l + \sum_{i=1}^{N-1} |d_i - \bar{d}|}{d_f + d_l + (N-1)\bar{d}} \quad (20)$$

where d_f and d_l are the Euclidean distances between the extreme solutions of the POF and the boundary solutions of the obtained nondominated set. d_i represents the Euclidean distance between consecutive solutions in the obtained nondominated set of solutions and \bar{d} is the average of all distances d_i , $i=1, 2, \dots, N-1$, assuming that there are N solutions on the best nondominated front. Note that the smaller the Δ Metric value, the better diversity of the obtained nondominated set.

5.3. Parameter settings

In this paper, the proposed MOEA/D-AAP is compared with four classical decomposition-based MOEAs, including MOEA/D-DE [12], MOEA/D-STM [22], MOEA/D-ACD [16], and MOEA/D-APS [35]. To make fair comparisons, MOEA/D-DE, MOEA/D-STM, and MOEA/D-ACD all adopt the PBI decomposition method with a fixed θ value. In MOEA/D-APS, the value of θ is linearly increased from 1 to 10 with generations as described in [35]. In MOEA/D-AAP, the value of θ is adaptively adjusted by the AAP strategy. The other parameters of these competitors are the same as recommended in [12], which are summarized as follows.

- 1) *Control parameters in DE and polynomial mutation*: $CR=0.5$ and $F=0.5$ in the DE operator; $p_m=1/n$ and $\eta=20$ in the polynomial mutation operator.
- 2) *Neighborhood size*: $T=20$.
- 3) *Probability to select in the neighborhood*: $\delta=0.9$.
- 4) *Control parameter in the replacement operator*: $n_r=2$.
- 5) *Population size*: $N=100$.
- 6) *Number of runs and stopping condition*: Each algorithm is run 30 times independently on each test instance. For each run, an algorithm stops after a given maximal number of generations, which was set to 100 for all the test instances.

6. Experimental results and discussion

6.1. F1-F6 test results

Table 3 gives the best, average, and worst values of IGD for the five competitors on F1-F6 test instances, in which the bold means the corresponding algorithm achieves the best results on the test instance. The differences between the approximations are assessed by the Wilcoxon rank-sum test at the 0.05 significance level. Signs of † and ‡ in the superscript form on mean values indicate the significance of the proposed method. For F1, F3-F5, compared with MOEA/D-DE and MOEA/D-STM, the IGD value of MOEA/D-APS and MOEA/D-AAP were both decreased significantly, which indicates that both APS and AAP strategies are capable of balancing convergence and diversity of the algorithm. Compared with MOEA/D-APS, the IGD value of MOEA/D-AAP is further reduced, which shows that the proposed AAP strategy is more beneficial to find boundary solutions and to improve the population diversity. For F2, it is easy for MOEA/D-DE, MOEA/D-STM and MOEA/D-APS to fall into a local POF, resulting in the high mean IGD value, while MOEA/D-ACD and MOEA/D-AAP has a stable performance and can approximate the whole POF in most cases. For the three-objective optimization problem F6, compared with MOEA/D-DE, MOEA/D-STM and MOEA/D-APS, the IGD values of MOEA/D-ACD and MOEA/D-AAP decrease to a certain extent. From Table 3, the AAP scheme can effectively improve the performance of the MOEA/D algorithm.

Table 3 IGD comparison results

| Prob. | | IGD | | | | |
|-------|-------|-----------|------------|-----------------|------------|---------------|
| | | MOEA/D-DE | MOEA/D-STM | MOEA/D-ACD | MOEA/D-APS | MOEA/D-AAP |
| F1 | Best | 0.1037 | 0.1033 | 0.0270 | 0.0601 | 0.0179 |
| | Mean | 0.1066 † | 0.1075 † | 0.0390 † | 0.0637 † | 0.0223 |
| | Worst | 0.1084 | 0.1140 | 0.0640 | 0.0673 | 0.0247 |
| F2 | Best | 0.0059 | 0.7838 | 0.0164 | 0.0073 | 0.0057 |
| | Mean | 0.4245 † | 0.7838 † | 0.0239 ‡ | 0.3818 † | 0.0613 |
| | Worst | 0.7838 | 0.7838 | 0.0399 | 0.7838 | 0.7838 |
| F3 | Best | 0.1959 | 0.1899 | 0.0278 | 0.1207 | 0.0195 |
| | Mean | 0.3582 † | 0.3918 † | 0.0410 † | 0.3110 † | 0.0242 |
| | Worst | 0.4014 | 0.4002 | 0.1763 | 0.3895 | 0.0548 |
| F4 | Best | 0.0811 | 0.0792 | 0.0167 | 0.0721 | 0.0161 |
| | Mean | 0.1283 † | 0.0834 † | 0.0208 † | 0.6167 † | 0.0170 |
| | Worst | 0.7525 | 0.0862 | 0.0274 | 0.7525 | 0.0181 |
| F5 | Best | 0.0212 | 0.0206 | 0.0085 | 0.0103 | 0.0078 |
| | Mean | 0.0217 † | 0.0216 † | 0.0173 † | 0.0191 † | 0.0096 |
| | Worst | 0.0225 | 0.0236 | 0.0265 | 0.0215 | 0.0127 |
| F6 | Best | 0.0531 | 0.0466 | 0.0358 | 0.0474 | 0.0389 |
| | Mean | 0.0587 † | 0.0503 † | 0.0409 | 0.0537 † | 0.0404 |
| | Worst | 0.0746 | 0.0543 | 0.0418 | 0.0617 | 0.0421 |

Wilcoxon's rank sum test at a 0.05 significance level is performed between MOEA/D-AAP and each of MOEA/D-DE, MOEA/D-STM, MOEA/D-ACD and MOEA/D-APS. † and ‡ denote the performance of the corresponding algorithm is significantly worse than or better than that of MOEA/D-AAP, respectively.

Table 4 gives the best, average, and worst values of HV for the five competitors on F1-F6 test instances, in which the bold means the corresponding algorithm achieves the best results on the test instance. It can be observed that, in terms of HV, MOEA/D-AAP significantly outperforms other compared algorithms, except for F2. MOEA/D-ACD achieves the best performance on F2.

Table 4 HV comparison results

| Prob. | | HV | | | | |
|-------|-------|-----------|------------|---------------|------------|---------------|
| | | MOEA/D-DE | MOEA/D-STM | MOEA/D-ACD | MOEA/D-APS | MOEA/D-AAP |
| F1 | Best | 1.0880 | 1.0886 | 1.1040 | 1.1007 | 1.1059 |
| | Mean | 1.0872 † | 1.0870 † | 1.1018 † | 1.0991 † | 1.1051 |
| | Worst | 1.0850 | 1.0845 | 1.1007 | 1.0970 | 1.1027 |
| F2 | Best | 0.3067 | 0.1100 | 0.3097 | 0.3046 | 0.3077 |

| | | | | | | |
|----|-------|---------------|----------|-----------------|----------|---------------|
| | Mean | 0.1938 † | 0.1100 † | 0.3063 ‡ | 0.1934 † | 0.2879 |
| | Worst | 0.1100 | 0.1100 | 0.2994 | 0.1100 | 0.1100 |
| F3 | Best | 1.1083 | 1.1081 | 1.1131 | 1.1154 | 1.1264 |
| | Mean | 1.0839 † | 1.0782 † | 1.1109 † | 1.0939 † | 1.1254 |
| | Worst | 1.0761 | 1.0753 | 1.0989 | 1.0822 | 1.1228 |
| F4 | Best | 1.0277 | 1.0279 | 1.0330 | 1.0291 | 1.0354 |
| | Mean | 0.9640 † | 1.0266 † | 1.0313 | 0.2934 † | 1.0344 |
| | Worst | 0.1100 | 1.0234 | 1.0284 | 0.1100 | 1.0317 |
| F5 | Best | 0.5276 | 0.5272 | 0.5248 | 0.5314 | 0.5327 |
| | Mean | 0.5263 † | 0.5257 † | 0.5197 † | 0.5260 † | 0.5310 |
| | Worst | 0.5250 | 0.5237 | 0.5120 | 0.5217 | 0.5242 |
| F6 | Best | 1.2540 | 1.2664 | 1.2713 | 1.2587 | 1.2724 |
| | Mean | 1.2459 † | 1.2624 | 1.2662 | 1.2509 † | 1.2698 |
| | Worst | 1.2263 | 1.2589 | 1.2569 | 1.2398 | 1.2658 |

Wilcoxon's rank sum test at a 0.05 significance level is performed between MOEA/D-AAP and each of the other competing algorithms. † and ‡ denote whether the performance of the corresponding algorithm is significantly worse than or better than that of MOEA/D-AAP, respectively.

Table 5 gives the best, average, and worst values of Δ Metric for the five competitors on F1-F6 test instances, in which the bold means the corresponding algorithm achieves the best results on the test instance. From Table 5, similar performance can be observed on the comparisons of five algorithms in terms of Δ Metric, where MOEA/D-AAP is significantly better than other compared algorithms on five out of six test problems. It should be noted that, for F1, although the Δ values of MOEA/D-DE are smaller than that of MOEA/D-AAP, it is evident from the observation in Fig. 6 that MOEA/D-DE cannot cover the edge region, and our MOEA/D-AAP can find more boundary solutions. Therefore, the AAP strategy designed in this paper can improve the coverage performance of MOEA/D in dealing with complicated MOPs.

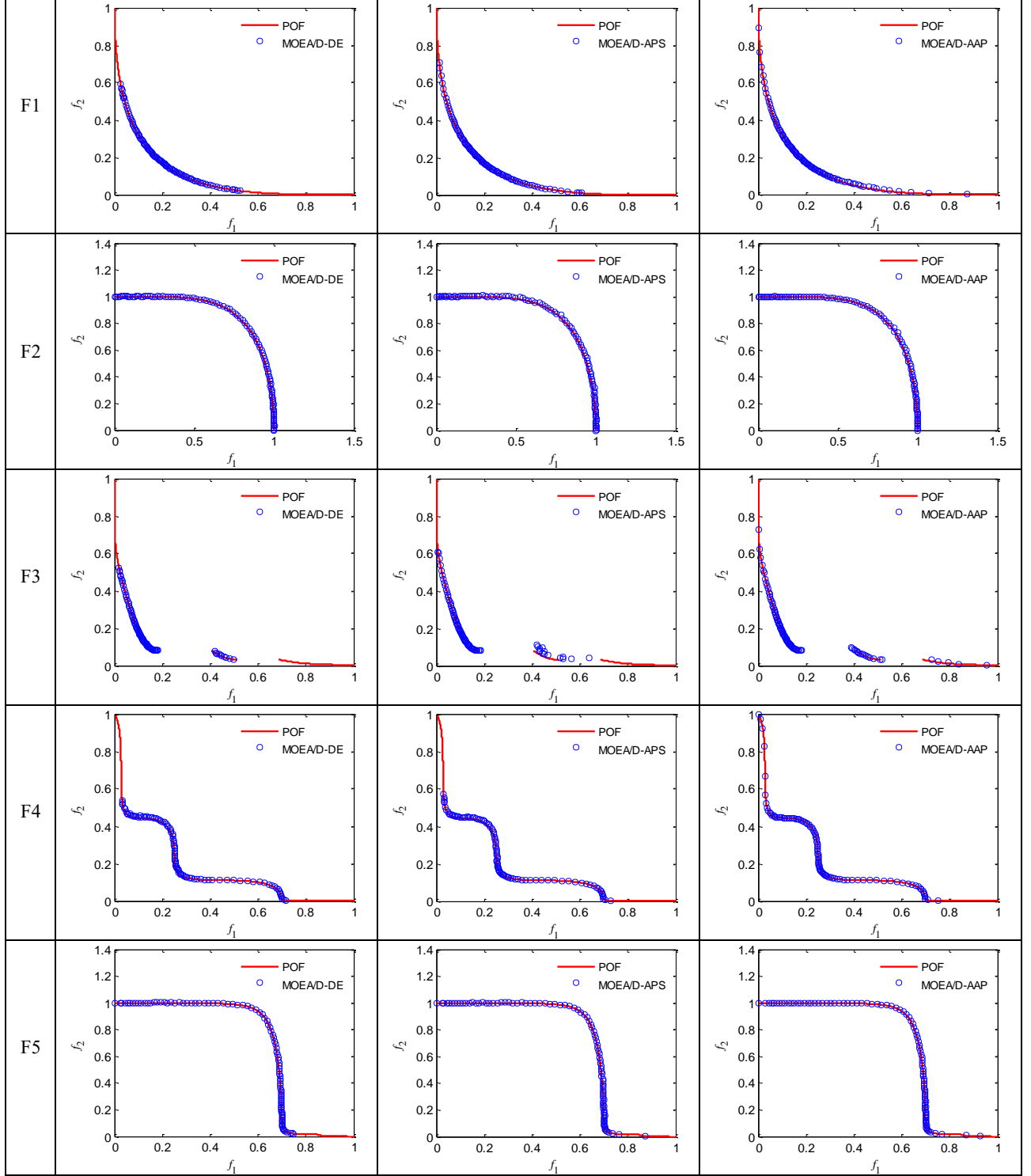
Table 5 Δ Metric comparison results

| Prob. | | Δ | | | | |
|-------|-------|-----------------|---------------|---------------|---------------|---------------|
| | | MOEA/D-DE | MOEA/D-STM | MOEA/D-ACD | MOEA/D-APS | MOEA/D-AAP |
| F1 | Best | 0.7040 | 0.7201 | 0.7274 | 0.7010 | 0.7949 |
| | Mean | 0.7187 ‡ | 0.7461 ‡ | 0.8161 | 0.7189 ‡ | 0.8114 |
| | Worst | 0.7398 | 0.7746 | 1.1455 | 0.7435 | 0.8312 |
| F2 | Best | 0.2173 | 1.0000 | 0.9195 | 0.2331 | 0.2447 |
| | Mean | 0.6895 † | 1.0000 † | 1.3617 † | 0.7436 † | 0.3212 |
| | Worst | 1.0541 | 1.0000 | 1.5676 | 1.3964 | 1.0000 |
| F3 | Best | 0.8643 | 0.9349 | 0.7238 | 0.8127 | 0.8478 |
| | Mean | 0.9193 | 0.9480 † | 1.1856 † | 0.9230 † | 0.9188 |
| | Worst | 0.9508 | 0.9592 | 1.6792 | 0.9527 | 0.9427 |
| F4 | Best | 0.6603 | 0.6784 | 0.5947 | 0.6751 | 0.6273 |
| | Mean | 0.6984 † | 0.7019 † | 0.7630 † | 0.9308 † | 0.6815 |
| | Worst | 1.0000 | 0.7242 | 1.1748 | 1.0000 | 0.7031 |
| F5 | Best | 0.4377 | 0.4281 | 0.3564 | 0.4172 | 0.3394 |
| | Mean | 0.4506 † | 0.4435 † | 0.4095 † | 0.4335 † | 0.3761 |
| | Worst | 0.4644 | 0.4756 | 0.6345 | 0.4587 | 0.4105 |
| F6 | Best | 0.2385 | 0.2197 | 0.2159 | 0.2235 | 0.2193 |
| | Mean | 0.2755 † | 0.2473 † | 0.2481 † | 0.2491 † | 0.2368 |
| | Worst | 0.3353 | 0.2867 | 0.2593 | 0.2659 | 0.2588 |

Wilcoxon's rank sum test at a 0.05 significance level is performed between MOEA/D-AAP and each of the other competing algorithms. † and ‡ denote whether the performance of the corresponding algorithm is significantly worse than or better than that of MOEA/D-AAP, respectively.

Fig. 6 shows the approximated POF for MOEA/D-DE, MOEA/D-APS, and MOEA/D-AAP when they obtain the lowest IGD values on F1 to F6. It can be seen from Fig. 6 that the solutions obtained by MOEA/D-DE and MOEA/D-APS converge to the middle region of the POF and it is difficult to find the boundary solutions, resulting in the

decrease of population diversity. The proposed MOEA/D-AAP is more likely to find the boundary solutions. Especially for F1, F3, and F4 test instances with convex POFs, Fig. 6 indicates that the proposed AAP scheme can obtain a better solution distribution along the actual POF. It is worth noting that, when the minimum IGD value is obtained, MOEA/D-DE, MOEA/D-APS, and MOEA/D-AAP achieve the same algorithm performance on F2 with a concave POF.



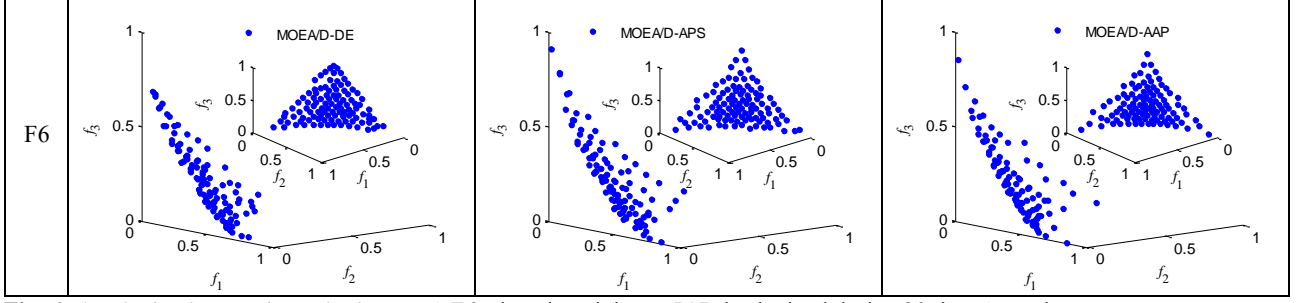


Fig. 6 The obtained approximated POF on F1-F6 when the minimum IGD is obtained during 30 times' running

Fig. 7 plots the evolution curves of the IGD metric value versus the number of generations for MOEA/D-DE, MOEA/D-STM, MOEA/D-APS and MOEA/D-AAP on each test instance when the minimum IGD is obtained during the 30 times' running. From Fig. 7, the proposed MOEA/D-AAP not only has the fastest convergence speed but also has the highest approximation accuracy on F1, F3, and F4. However, the IGD value of other three algorithms can easily stagnate due to the loss of population diversity. For F2, the final approximation accuracy of MOEA/D-DE, MOEA/D-APS and MOEA/D-AAP is about the same, whereas MOEA/D-AAP has the fastest convergence speed. For F5 and F6, the AAP scheme can offer a slight improvement on the IGD metric.

In order to more clearly illustrate the proposed AAP strategy, we also present boxplots of the IGD and HV metrics obtained by MOEA/D-DE, MOEA/D-STM, MOEA/D-APS, and MOEA/D-AAP from 30 independent runs on each instance in Figs. 8 and 9, respectively. These results clearly show that MOEA/D-AAP is the best. And, more remarkably, for F2, the solutions obtained by MOEA/D-DE, MOEA/D-STM and MOEA/D-APS are easy to converge to the peak point of the POF, while the solutions obtained by the proposed MOEA/D-AAP can cover the entire POF in most cases. Thus, we can conclude that the use of the AAP scheme has successfully improved the algorithm performance on MOPs with complicated POFs.

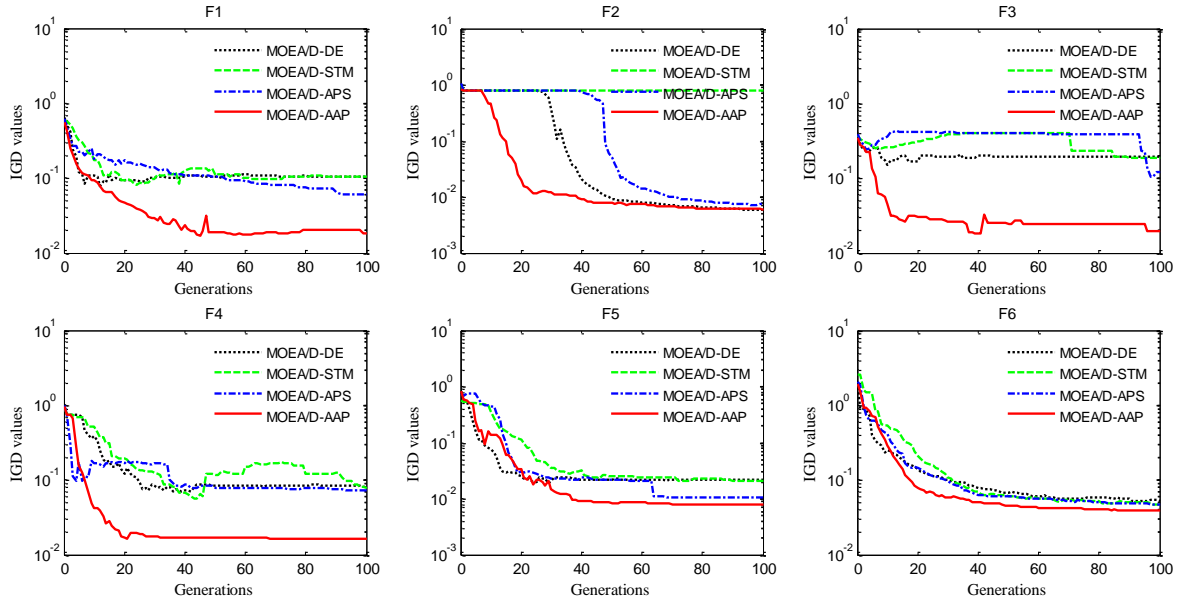


Fig. 7 The IGD evolution curves of F1-F6 when the minimum IGD value is obtained during the 30 times' running

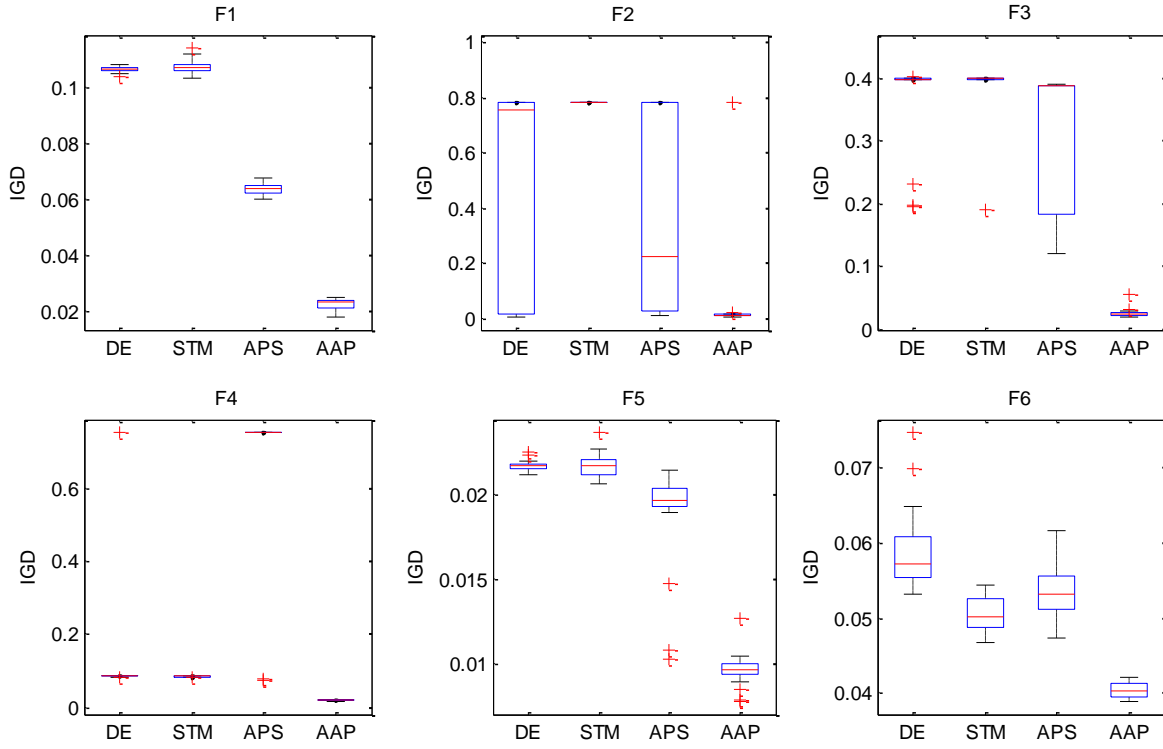


Fig. 8 Boxplots for the comparison of MOEA/D-AAP with MOEA/D-DE, MOEA/D-STM and MOEA/D-APS on IGD

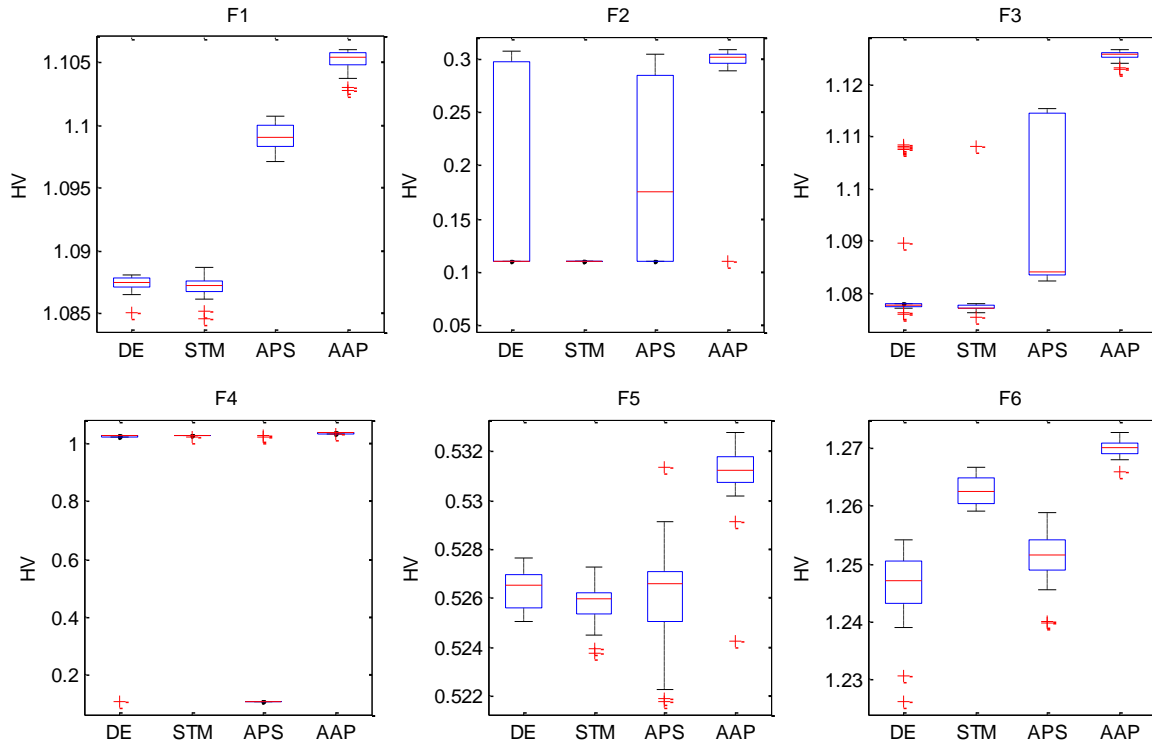


Fig. 9 Boxplots for the comparison of MOEA/D-AAP with MOEA/D-DE, MOEA/D-STM and MOEA/D-APS on HV

6.2. Determination of the range of penalty factor

In this study, the range of θ value is set to $[2, 50]$. Let us briefly analyze how to determine the range. First, we clarify that it is proper for the upper and lower limit of θ value to be integer. Then, we analyze the setting of the lower limit of the θ value. When θ is set to 0, the vertical distance d_2 , which represents the diversity, does not work

when calculating the aggregation function value for each subproblem. Therefore, the case of $\theta = 0$ should be removed. When θ is set to 1, the formula $\theta(con)=1/(1+ae^{-b \times con})$ cannot be used to compute an effective value for coefficient a . Hence, the case of $\theta = 1$ should be removed. In conclusion, the lower limit of the θ value is set to 2. The upper limit of θ value is set to $\zeta \times N$, where N is the population size and ζ is a positive value. In MOEA/D-AAP, the value of ζ which can determine the magnitude of the penalty factor for each subproblem needs to be studied carefully. In this subsection, we examine the impact of ζ on the AAP strategy using F1 and F2 test problems. In the experiments, the value of ζ varies from 0.1 to 0.9. The remaining parameters of the MOEA/D-AAP are set to the same as in Section 5.3.

Table 5 gives the obtained IGD values of MOEA/D-AAP with different ζ values on F1 and F2. From Table 5, for the convex optimization problem F1, the best, mean, and worst values of IGD are better than others when $\zeta = 0.5$. For the concave optimization problem F2, the value of ζ has little effect on the algorithm's performance in terms of the best IGD values except $\zeta = 0.9$. It is worth noting that the value of ζ needs to be considered carefully for other test problems, although $\zeta = 0.5$ can significantly improve the algorithm's performance on the six benchmark test problems in this paper.

Table 5 IGD comparison results with different ζ values

| Prob. | | IGD | | | | |
|-------|-------|---------------|---------------|---------------|---------------|---------------|
| | | $\zeta = 0.1$ | $\zeta = 0.3$ | $\zeta = 0.5$ | $\zeta = 0.7$ | $\zeta = 0.9$ |
| F1 | Best | 0.0608 | 0.0246 | 0.0179 | 0.0182 | 0.0205 |
| | mean | 0.0662 | 0.0297 | 0.0223 | 0.0236 | 0.0229 |
| | worst | 0.0698 | 0.0326 | 0.0247 | 0.0303 | 0.0305 |
| F2 | Best | 0.0057 | 0.0059 | 0.0057 | 0.0059 | 0.0108 |
| | mean | 0.5268 | 0.0607 | 0.0613 | 0.0615 | 0.0255 |
| | worst | 0.7838 | 0.7838 | 0.7838 | 0.7838 | 0.0283 |

6.3. Sensitivity analysis

In this section, we investigate the sensitivity of the control parameters in MOEA/D-AAP, including the control parameter ζ in the AAP scheme, the probability δ for the mating restriction, and the control parameters F and CR in the DE operator, on both F1 and F3. As the control parameter ζ and δ play an interdependent role on the performance of MOEA/D-AAP, we set combinations of different values of ζ and δ (e. g., 0.1, 0.3, 0.5, 0.7 and 0.9) in the experiments. The remaining parameters are the same as those in Section 5.3. Fig. 10 shows the mean IGD values of the approximated POFs obtained by MOEA/D-AAP with different control parameter ζ and δ on F1 and F3. It should be noted that the POF of the test function F3 not only has a sharp peak and a long tail, but also has disconnected regions, which is very suitable for testing the performance of MOEA/D-AAP. Fig. 10 shows that MOEA/D-AAP with different ζ and δ values may have the different performance. MOEA/D-AAP is very sensitive to ζ and a suitable ζ value is required for satisfactory performance.

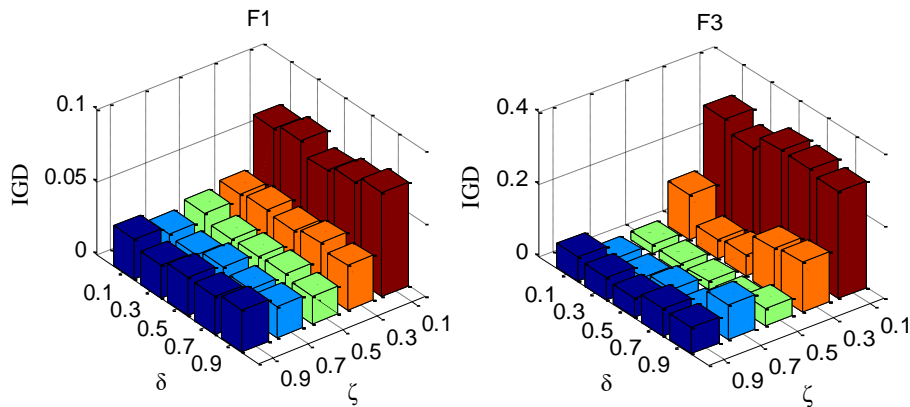


Fig. 10 The mean IGD values obtained by MOEA/D-AAP with different combinations of ζ and δ on F1 and F3 over 30 runs

To investigate the sensitivity of the parameters F and CR in the DE operator, we set combination of five F values (0.1, 0.3, 0.5, 0.7 and 0.9) and five CR values (0.1, 0.3, 0.5, 0.7 and 0.9) on F1 and F3 in the experiments. The other parameters are the same as those in Section 5.3. Fig. 11 shows the mean IGD values of the approximated POFs achieved by MOEA/D-AAP with different combinations of F and CR on F1 and F3 over 30 runs. It can be observed

that when F is set to be a value near 0.5 and CR is set to be a relatively large values (e. g., 0.5, 0.7), MOEA/D-AAP achieves better performance.

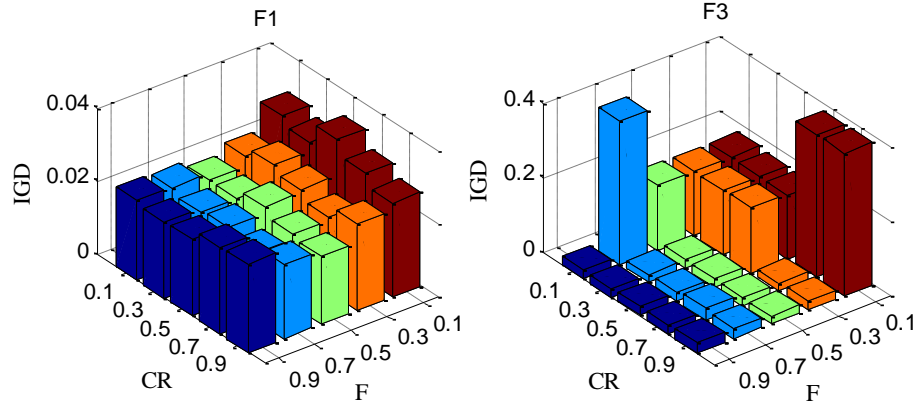


Fig. 11 The mean IGD values obtained by MOEA/D-AAP with different combinations of F and CR on F1 and F3 over 30 runs

6.4. Multiobjective optimization of the wastewater treatment process

The aim of the multiobjective optimization in the WWTP is to dynamically optimize the set points of dissolved oxygen concentration (S_O) in the aerobic zone and nitrate nitrogen concentration (S_{NO}) in the anaerobic zone in order to reduce energy consumption (EC) and amercement (AM). Based on multiobjective optimization control, the treatment effect can be improved and the operation cost can be reduced [36]. In this section, the proposed MOEA/D-AAP is adopted to solve the MOP in the WWTP based on the BSM1 platform developed by International Water Association (IWA) and European Cooperation in Science and Technology (COST). The schematic representation of the WWTP is shown in Fig. 12. Based on the analysis of the reaction mechanism of the WWTP, the wastage flow rate Q_w , the oxygen transfer coefficient K_{La3} , K_{La4} , the $S_{O,5}$, and $S_{NO,2}$ are selected as the decision variables (or input variables). Table 6 gives the range of the value of each decision variable.

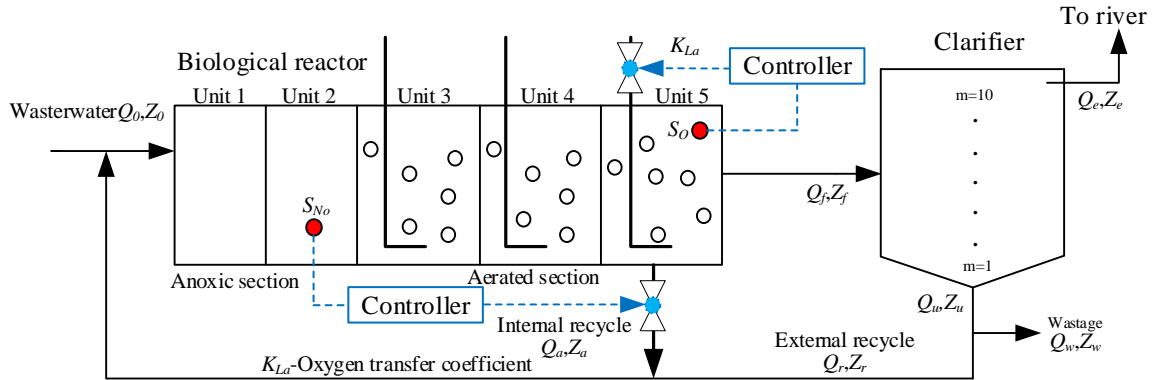


Fig. 12 Schematic representation of the wastewater treatment plant

Table 6 The range of value for decision variables

| Decision variables | Default values | Lower limit | Upper limit |
|--|----------------|-------------|-------------|
| $Q_w/\text{m}^3 \cdot \text{d}^{-1}$ | 385 | 100 | 800 |
| K_{La3}/d^{-1} | 240 | 0 | 240 |
| K_{La4}/d^{-1} | 240 | 0 | 240 |
| $S_{O,5}/\text{mg} \cdot \text{L}^{-1}$ | 2 | 0.4 | 3 |
| $S_{NO,2}/\text{mg} \cdot \text{L}^{-1}$ | 1 | 0.2 | 2 |

The MOP in the WWTP is described as follows [31]:

$$\begin{aligned} \min \mathbf{F}(\mathbf{x}) &= (f_{\text{EC}}(\mathbf{x}), f_{\text{AM}}(\mathbf{x}))^T \\ \text{s.t. } &\begin{cases} g_j(\mathbf{x}) \leq c_j, j=1,2,\dots,5 \\ x_i^l \leq x_i \leq x_i^u, i=1,2,\dots,5 \end{cases} \end{aligned} \quad (21)$$

where $\mathbf{x}=(x_1, x_2, x_3, x_4, x_5)^T=[Q_w, K_{La3}, K_{La4}, S_{O,5}, S_{NO,2}]^T$ are five decision variables, $f_{\text{EC}}(\mathbf{x})$ and $f_{\text{AM}}(\mathbf{x})$ denote the objective functions of EC and AM, respectively; $g_j(\mathbf{x})$ is the j th inequality constraint which represents the relationship between the water quality parameter and the decision vectors, $j=1,2,\dots,5$; x_i^l and x_i^u are the lower and upper bounds of the i th decision variable, respectively.

The procedure of multiobjective optimization in the WWTP is as follows: Firstly, to establish the nonlinear relationship between the optimization objectives and the decision variables, a data-driven modeling method is used to design the model of EC and AM [36]. Secondly, the proposed MOEA/D-AAP in this study is adopted to deal with the MOP in the decision space to find a set of Pareto solutions. In addition, an intelligent decision-maker (IDM) which is applied to select a compromise solution from the POS is developed by a fuzzy membership function approach [7]. Finally, the optimized set points of S_O and S_{NO} are controlled by an intelligent multivariable controller [38].

The satisfactory degree of Pareto solution \mathbf{x}_k for the i th objective function f_i is expressed by a membership function, defined as follows:

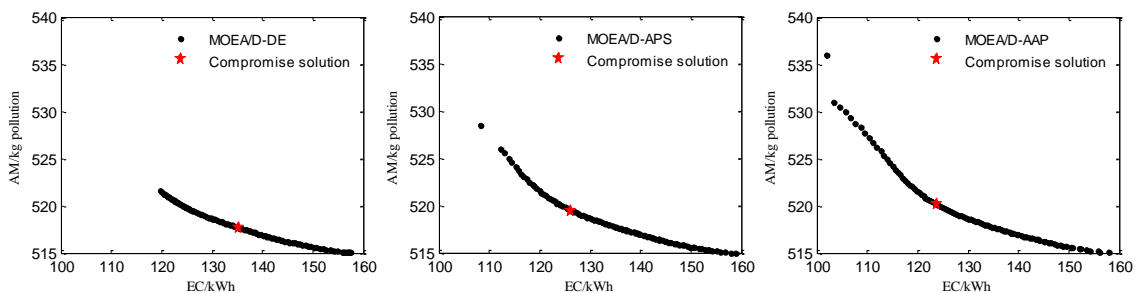
$$\mu_i^k = \begin{cases} 1, & f_i(\mathbf{x}_k) \leq f_i^{\min} \\ \frac{f_i^{\max} - f_i(\mathbf{x}_k)}{f_i^{\max} - f_i^{\min}}, & f_i^{\min} < f_i(\mathbf{x}_k) < f_i^{\max} \\ 0, & f_i(\mathbf{x}_k) \geq f_i^{\max} \end{cases} \quad (22)$$

where f_i^{\max} and f_i^{\min} are the maximum and minimum of the i th objective function f_i , respectively. Then, the normalized membership function μ^k of \mathbf{x}_k is calculated by

$$\mu^k = \frac{\sum_{i=1}^m \mu_i^k}{\sum_{k=1}^{|S|} \sum_{i=1}^m \mu_i^k} \quad (23)$$

where m is the number of objective functions, $|S|$ is the element number of POS obtained by the algorithm. In this study, the compromise solution selected by IDM is the one having the maximum of μ^k in the POS.

Fig. 11 shows the approximated POFs and HV evolution curves obtained by MOEA/D-DE, MOEA/D-APS, and MOEA/D-AAP. It can be seen from Fig. 13 that the diversity of MOEA/D-AAP is better than MOEA/D-DE and MOEA/D-APS. Particularly, the AAP scheme can find some boundary solutions on the MOP in the WWTP. From the plot of HV curves, MOEA/D-AAP achieves a higher HV value compared with MOEA/D-DE and MOEA/D-APS. It is noteworthy that normalization algorithm is needed to be performed when calculating the value of HV. Table 7 gives the optimization results of EC and AM in the WWTP. Some conclusions can be drawn by analyzing results of Table 7: 1) **Both** EC and AM can be reduced by optimizing the set points in the WWTP; 2) Different penalty factors in the PBI decomposition method can lead to different performance of MOEA/D; 3) The proposed AAP scheme can enhance the performance of MOEA/D for solving the MOP in the WWTP; 4) AM rises with the reduction of EC, which indicates that EC and AM are two conflict objectives.



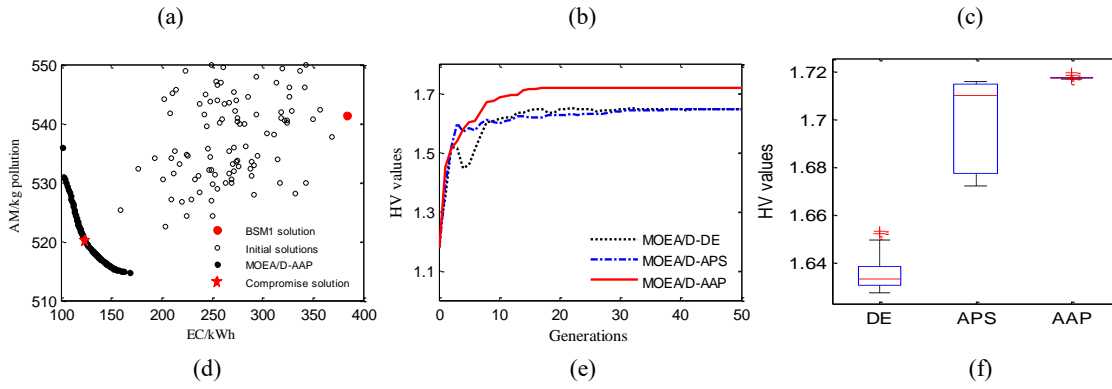


Fig. 13 Approximated POFs and HV evolution curves

(a) MOEA/D-DE with $\theta=2$; (b) MOEA/D-APS with $\theta=1\sim5$; (c) MOEA/D-AAP;
(d) Optimization process; (e) HV evolution curves; (f) Boxplots on HV

Table 7 Multiobjective optimization results in the WWTP

| Method | θ | EC (kWh) | Up/Down | AM (kg poll.) | Up/Down |
|------------|------------------|---------------|----------------|---------------|---------------|
| Default | — | 383.73 | — | 541.26 | — |
| MOEA/D-DE | $\theta=2$ | 135.19 | ↓64.77% | 517.67 | ↓4.36% |
| | $\theta=5$ | 125.30 | ↓67.34% | 519.68 | ↓3.99% |
| MOEA/D-STM | $\theta=2$ | 132.25 | ↓65.54% | 517.12 | ↓4.46% |
| | $\theta=5$ | 130.74 | ↓65.93% | 519.45 | ↓4.03% |
| MOEA/D-ACD | $\theta=2$ | 130.80 | ↓65.91% | 519.36 | ↓4.05% |
| | $\theta=5$ | 126.77 | ↓66.96% | 519.64 | ↓3.99% |
| MOEA/D-APS | $\theta=1\sim5$ | 126.18 | ↓67.11% | 519.45 | ↓4.03% |
| | $\theta=1\sim10$ | 125.79 | ↓67.21% | 519.55 | ↓4.01% |
| MOEA/D-AAP | adaptive | 123.69 | ↓67.77% | 520.14 | ↓3.90% |

Note: Default represents the results obtained by BSM1 with default settings.

7. Conclusion

Since the traditional MOEA/D algorithm which uses the PBI method with a fixed penalty factor cannot cover the whole POF in solving MOPs with complicated POFs, an angle-based adaptive penalty scheme is developed to dynamically adjust the penalty factor of each subproblem. In the evolutionary process, the solution concentration of each weight vector is calculated by using the acute angle between solutions and weight vectors. Combining with a sigmoid mapping function, the AAP scheme is used to tradeoff the convergence and diversity. The performance of the proposed MOEA/D-AAP is verified on six benchmark test instances with complicated POFs and the MOP in the WWTP. The experimental results demonstrate that the proposed AAP scheme can enhance the ability of MOEA/D to find boundary solutions. Since the replacement neighborhood size is crucial to the performance of MOEA/D, the next work is to design an MOEA/D with adaptive replacement scheme, and apply it to the whole process optimization in the WWTP.

Acknowledgment

This work was supported in part by the National Science Foundation for Distinguished Young Scholars of China (Grant No. 61225016) and the State Key Program of National Natural Science of China (Grant No. 61533002).

References

- [1] A. Trivedi, D. Srinivasan, K. Sanyal, G. Abhiroop. A survey of multiobjective evolutionary algorithms based on decomposition. IEEE Transactions on Evolutionary Computation, 2017, 21(3): 440-462.

- [2] I. Giagkiozis, P. J. Fleming. Methods for multi-objective optimization: An analysis. *Information Sciences*, 2015, 293: 338-350.
- [3] A. M. Zhou, Q. F. Zhang. Are all the subproblems equally important? Resource allocation in decomposition-based multiobjective evolutionary algorithms. *IEEE Transactions on Evolutionary Computation*, 2016, 20(1): 52-64.
- [4] K. Li, S. Kwong, J. J. Cao, M. Q. Li, J. H. Zheng, R. M. Shen. Achieving balance between proximity and diversity in multi-objective evolutionary algorithm. *Information Sciences*, 2012, 182(1): 220-242.
- [5] C. A. C. Coello, G. T. Pulido, M. S. Lechuga. Handling multiple objectives with particle swarm optimization. *IEEE Transactions on Evolutionary Computation*, 2004, 8(3): 256-279.
- [6] K. Deb, A. Pratap, S. Agarwal, T. Meyarivan. A fast and elitist multiobjective genetic algorithm: NSGA-II. *IEEE Transactions on Evolutionary Computation*, 2002, 6(2): 182-197.
- [7] E. Zitzler, M. Laumanns, L. Thiele. SPEA2: Improving the strength Pareto evolutionary algorithm. 2001.
- [8] N. Beume, B. Naujoks, M. Emmerich. SMS-EMOA: Multiobjective selection based on dominated hypervolume. *European Journal of Operational Research*, 2007, 181(3): 1653-1669.
- [9] J. Bader, E. Zitzler. HypE: An algorithm for fast hypervolume-based many-objective optimization. *Evolutionary Computation*, 2011, 19(1): 45-76.
- [10] A. Jaszkiewicz. On the performance of multiple-objective genetic local search on the 0/1 knapsack problem-a comparative experiment. *IEEE Transactions on Evolutionary Computation*, 2002, 6(4): 402-412.
- [11] Q. F. Zhang, H. Li. MOEA/D: A multiobjective evolutionary algorithm based on decomposition. *IEEE Transactions on Evolutionary Computation*, 2007, 11(6): 712-731.
- [12] H. Li, Q. F. Zhang. Multiobjective optimization problems with complicated Pareto sets, MOEA/D and NSGA-II. *IEEE Transactions on Evolutionary Computation*, 2009, 13(2): 284-302.
- [13] Y. T. Qi, X. L. Ma, F. Liu, L. C. Jiao, J. Y. Sun, J. S. Wu. MOEA/D with adaptive weight adjustment. *Evolutionary Computation*, 2014, 22(2): 231-264.
- [14] K. Li, R. Z. Chen, G. Y. Min, X. Yao. Integration of preferences in decomposition multiobjective optimization. *IEEE Transactions on Cybernetics*, 2018.
- [15] H. L. Liu, F. Q. Gu, Q. F. Zhang. Decomposition of a multiobjective optimization problem into a number of simple multiobjective subproblems. *IEEE Transactions Evolutionary Computation*, 2014, 18(3): 450-455.
- [16] L. P. Wang, Q. F. Zhang, A. M. Zhou, M. G. Gong, L. C. Jiao. Constrained Subproblems in a Decomposition-Based Multiobjective Evolutionary Algorithm. *IEEE Transactions on Evolutionary Computation*, 2016, 20(3): 475-480.
- [17] X. Y. Cai, Z. W. Mei, Z. Fan, Q. F. Zhang. A constrained decomposition approach with grids for evolutionary multiobjective optimization. *IEEE Transactions on Evolutionary Computation*, 2017.
- [18] M. Y. Wu, K. Li, S. Kwong, Q. F. Zhang, J. Zhang. Learning to decompose: a paradigm for decomposition-based multiobjective optimization. *IEEE Transactions on Evolutionary Computation*, 2018.
- [19] X. Y. Cai, Y. X. Li, Z. Fan, Q. F. Zhang. An external archive guided multiobjective evolutionary algorithm based on decomposition for combinatorial optimization. *IEEE Transactions on Evolutionary Computation*, 2015, 19(4): 508-523.
- [20] X. Y. Cai, H. R. Sun, Q. F. Zhang, Y. H. Huang. A grid weighted sum Pareto local search for combinatorial multi and many-objective optimization. *IEEE transactions on cybernetics*, 2018.
- [21] K. Li, Á. Fialho, S. Kwong, Q. Zhang. Adaptive operator Selection with bandits for a multiobjective evolutionary algorithm based on decomposition. *IEEE Transactions on Evolutionary Computation*, 2014, 18(1): 114-130.
- [22] K. Li, Q. F. Zhang, S. Kwong, M. Q. Li, R. Wang. Stable matching-based selection in evolutionary multiobjective optimization. *IEEE Transactions on Evolutionary Computation*, 2014, 18(6): 909-923.
- [23] K. Li, S. Kwong, Q. F. Zhang, K. Deb. Interrelationship-based selection for decomposition multiobjective optimization. *IEEE Transactions on Cybernetics*, 2015, 45(10): 2076-2088.
- [24] Z. K. Wang, Q. F. Zhang, A. M. Zhou, M. G. Gong, L. C. Jiao. Adaptive replacement strategies for MOEA/D. *IEEE Transactions on cybernetics*, 2016, 46(2): 474-486.
- [25] K. Li, K. Deb, Q. F. Zhang, S. Kwong. An evolutionary many-objective optimization algorithm based on dominance and decomposition. *IEEE Transactions on Evolutionary Computation*, 2015, 19(5): 694-716.
- [26] X. Y. Cai, Z. W. Mei, Z. Fan. A decomposition-based many-objective evolutionary algorithm with two types of adjustments for direction vectors. *IEEE Transactions on Cybernetics*, 2017.
- [27] X. Y. Cai, H. R. Sun, Z. Fan. A diversity indicator based on reference vectors for many-objective optimization. *Information Sciences*, 2018, 430: 467-486.
- [28] R. Z. Chen, K. Li, X. Yao. Dynamic multiobjectives optimization with a changing number of objectives. *IEEE Transactions on Evolutionary Computation*, 2018, 22(1): 157-171.
- [29] M. A. Jan, R. A. Khanum. A study of two penalty-parameterless constraint handling techniques in the framework of MOEA/D. *Applied Soft Computing*, 2013, 13(1): 128-148.
- [30] K. Li, R. Z. Chen, G. T. Fu, X. Yao. Two-archive evolutionary algorithm for constrained multi-objective optimization. *IEEE Transactions on Evolutionary Computation*, 2018.
- [31] X. Y. Cai, Z. X. Yang, Z. Fan, Q. Zhang. Decomposition-based-sorting and angle-based-selection for evolutionary multiobjective and many-objective optimization. *IEEE Transactions on Cybernetics*, 2017, 47(9): 2824-2837.

- [32] M. Y. Wu, K. Li, S. Kwong, Y. Zhou, Q. Zhang. Matching-based selection with incomplete lists for decomposition multi-objective optimization. *IEEE Transactions on Evolutionary Computation*, 2017, 21(4): 554-568.
- [33] S. Jiang, S. Yang. An improved multiobjective optimization evolutionary algorithm based on decomposition for complex Pareto fronts. *IEEE Transactions on Cybernetics*, 2016, 46(2): 421-437.
- [34] H. Sato. Inverted PBI in MOEA/D and its impact on the search performance on multi and many-objective optimization[C]//*Proceedings of the 2014 Annual Conference on Genetic and Evolutionary Computation*. ACM, 2014: 645-652.
- [35] S. Yang, S. Jiang, Y. Jiang. Improving the multiobjective evolutionary algorithm based on decomposition with new penalty schemes. *Soft Computing*, 2017, 21(16): 4677-4691.
- [36] J. F. Qiao, H. B. Zhou. Prediction of effluent total phosphorus based on self-organizing fuzzy neural network. *Control Theory and Applications*, 2017, 34(2): 224-232.
- [37] H. B. Zhou, J. F. Qiao. Optimal control of wastewater treatment process using hybrid multi-objective barebones particle swarm optimization algorithm. *CIESC Journal*, 2017, 68(9): 3511-3521.
- [38] H. B. Zhou. Dissolved oxygen control of the wastewater treatment process using self-organizing fuzzy neural network. *CIESC Journal*, 2017, 68(4): 1516-1524.